

APPLICATION

FOR

UNITED STATES LETTERS PATENT

APPLICANT NAME: Moon et al.

TITLE: MICR-BASED OPTICAL CHARACTER RECOGNITION
SYSTEM AND METHOD

DOCKET NO.: CHA920010020US1

INTERNATIONAL BUSINESS MACHINES CORPORATION

CERTIFICATE OF MAILING UNDER 37 CFR 1.10

I hereby certify that, on the date shown below, this correspondence is being deposited with the United States Postal Service in an envelope addressed to the Commissioner for Patents, Box Patent Application, Washington, D.C. 20231 as "Express Mail Patent to Addressee" Mailing Label No. EV0558486521 US

Mailing Label No. EV055848652US

on February 22, 2002

Wendy E. Thompson

Name of person mailing paper

Signature Wendy E. Thompson Date 2/22/02

Signature _____

Date _____

MICR-BASED OPTICAL CHARACTER RECOGNITION SYSTEM AND METHOD

BACKGROUND OF THE INVENTION

1. Technical Field

5 The present invention relates generally to character recognition systems, and relates more particularly to an optical character recognition system for reading E13B character data utilizing a MICR-based algorithm.

2. Related Art

10 Various applications require the automated processing of materials having printed characters. One common application involves the processing of checks by, e.g., a bank, a cashier at a retail store, etc. A popular technology for processing checks reads a set of special characters at the bottom of the check, which are called Magnetic Ink Character Recognition (MICR) characters. These characters are generally printed in a font referred to as E13B, and contain information regarding account number, routing numbers, etc. To
15 implement such a system, the characters are printed in a magnetic ink so that the characters can be read magnetically. Accordingly, MICR technology is limited to applications that utilize magnetic ink (e.g., checks printed on home computers that do not use magnetic ink cannot be read by MICR systems).

Once the characters are magnetically read, well established MICR-based algorithms are implemented to identify each E13B character. MICR-based systems and algorithms are described, for example, in U.S. Patent 6,243,504 B1, "Integrated Magnetic Ink Character Recognition System and Method Therefor," issued on June 5, 2001 to Kruppa, and U.S. Patent 5,026,974, Method for Recognizing the Leading Edge of a Character in E13B Font," issued on June 25, 1991 to Franlin et al., which are hereby incorporated by reference.

While most MICR systems utilize a single gap read head that provides marginal read-rate performance, IBM Corporation's 3890 system utilizes a multigap read head, along with a multigap MICR algorithm, to provide higher performance. Although multigap MICR systems generally provide superior read rates (i.e., low error rate), the systems are relatively expensive since they require multigap magnetic read heads.

A potential solution to the aforementioned problems is to use multiple lower cost recognition engines that perform independent analysis of the characters being read (referred to herein as "a multi-voting character recognition system"). A relatively inexpensive recognition engine may be implemented using known OCR (optical character reader) technology. OCR technology reads characters electro-optically and therefore does not require magnetic ink.

However, in order to provide a high level of accuracy, each character recognition engine must implement a unique feature set or feature vector for identifying characters. Unfortunately, the costs of developing, testing, and implementing new feature sets may be

significant. Accordingly, a need exists for independent feature sets for recognizing characters in a multi-voting recognition system that can utilize low cost technologies, such as OCR.

SUMMARY OF THE INVENTION

The present invention addresses the above mentioned problems, as well as others, by providing an OCR system that utilizes a MICR-based algorithm to identify E13B characters. In a first aspect, the invention provides a character recognition system, comprising: an optical character reader system for collecting character data by electro-optically scanning printed characters; a conversion system for converting the character data to a Magnetic Ink Character Recognition (MICR) format; and a recognition engine for interpreting the converted character data using a MICR algorithm.

In a second aspect, the invention provides a method for performing character recognition, comprising: collecting character data by electro-optically scanning printed characters; converting the character data to a Magnetic Ink Character Recognition (MICR) format; and interpreting the converted character data using a MICR algorithm.

In a third aspect, the invention provides a program product stored on a recordable medium for performing character recognition, comprising: means for accessing character data collected by an electro-optical scanning system; means for converting the character data to a Magnetic Ink Character Recognition (MICR) format; and means for interpreting the converted character data using a MICR algorithm.

10

BRIEF DESCRIPTION OF THE DRAWINGS

15

Figure 1 depicts a block diagram of a MICR-based OCR system in accordance with the present invention.

Figure 2 depicts a block diagram of a multi-voting character recognition engine in accordance with the present invention.

DETAILED DESCRIPTION OF THE INVENTION

Referring now to the drawings, Figure 1 depicts a character recognition system 10 that outputs a set of recognized characters 30 based on an inputted document 12 having character data printed in an E13B font. As noted above, E13B is commonly used on checks and the like to identify account numbers, routing information, check numbers, etc. As described below, character recognition system 10 utilizes a MICR-based OCR technology that provides a unique feature set for recognizing E13B character data.

Character recognition system 10 comprises an OCR scanner 14, a conversion system 18, and a MICR recognition engine 26. OCR scanner 14 may utilize any well-known electro-optical scanning technology to read in E13B character data on document 12 and generate scanned character data 16. The scanned character data may be saved at any density and using any format. In this exemplary embodiment, the scanned character data 16 is scanned by OCR scanner 14 in a grey-scale format at a pel (pixel element) density of between 200 and 600 dpi (dots per inch), which is typical for today's electro-optical scanner technology.

Conversion system 18 converts the scanned character data 16 into a MICR format character data 24. Specifically, the MICR format character data 24 has the equivalent spatial resolution and density as if captured by an actual multigap MICR read head. To achieve this, conversion system 18 comprises a scaling system 20, and a black and white format system 22. Scaling system 20 scales the resolution of the scanned character data to approximately 0.33 millimeters/pixel in the horizontal dimension and 0.43

millimeters/pixel in the vertical dimension. Scaling system 20 may utilize one or more techniques for changing the resolution (and therefore the size) of an image. Scaling techniques are well known in the art, and include, for example, U.S. Patent Application Serial Number _____, entitled "IMAGE SIZE REDUCTION METHOD AND SYSTEM," filed on January 31, 2002 by Prakash, which is hereby incorporated by reference.

In addition to the scaling operation performed by scaling system 20, black and white format system 22 converts scanned character data 16 from a grey-scale format to a black and white format. The result is a set of MICR format character data 24 that has the equivalent specifications of data obtained by a multigap MICR read head.

The MICR format character data 24 is then inputted to a MICR recognition engine 26 that analyzes the data and outputs a set of recognized characters 30. MICR recognition engine 26 may include any known MICR algorithm 28 for interpreting the MICR format character data 24. As noted above, such MICR algorithms are well known in the art and therefore will not be described in further detail herein. Accordingly, the present embodiment utilizes standard OCR scanning techniques in combination with well known MICR algorithms to provide a unique feature set for recognizing E13B data.

It should be recognized that while the MICR format character data 24 of the present embodiment comprises specifications based on a multigap MICR read head, the MICR format character data 24 could be based on specifications of a single gap MICR read head.

Referring now to Figure 2, a multi-voting character recognition system 50 is shown that utilizes multiple recognition systems 10, 42, 44. In this case, each of the recognition systems reads in E13B character data from document 40. One of the recognition systems comprises a MICR-based OCR recognition system 10, as described in Figure 1. The remaining recognition systems 42, 44 may utilize any other known engines/feature sets for identifying E13B character data. A voting system 46 obtains the results from each system 10, 42, 44, and outputs a final set of recognized characters 48 based on some voting logic implemented by voting system 46. For instance, if two of the three systems identifies a character as an "8" and the other system identifies the same character as a "3," then the voting system may determine that the character was indeed an "8" based on a 2-1 vote. Obviously, any type of voting scheme may be utilized, and any number of recognition systems may be implemented without departing from the scope of the invention.

It is understood that the present invention can be realized in hardware, software, or a combination of hardware and software. Moreover, any necessary computer systems for implementing the present invention can be realized in a centralized fashion in a single computerized workstation, or in a distributed fashion where different elements are spread across several interconnected systems (e.g., a network). Any kind of computer/server system(s) - or other apparatus adapted for carrying out the methods described herein - is suited. A typical combination of hardware and software could be a general purpose computer system with a computer program that, when loaded and executed, controls a

computer system such that it carries out the methods described herein. Alternatively, a specific use computer, containing specialized hardware for carrying out one or more of the functional tasks of the invention could be utilized. The present invention can also be embedded in a computer program product, which comprises all the features enabling the implementation of the methods described herein, and which - when loaded in a computer system - is able to carry out these methods. Computer program, software program, program, or software, in the present context mean any expression, in any language, code or notation, of a set of instructions intended to cause a system having an information processing capability to perform a particular function either directly or after either or both of the following: (a) conversion to another language, code or notation; and/or (b) reproduction in a different material form.

The foregoing description of the embodiments of this invention has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed, and obviously, many modifications and variations are possible. Such modifications and variations that may be apparent to a person skilled in the art are intended to be included within the scope of this invention as defined by the accompanying claims.